here the subscripted functions are polynomials. To apply the general equation solving chnique given at the beginning of section 2 to this problem, degree bounds must be stablished on the polynomials.

The multivariate problem is even more difficult than the rational function problem. lere we are given a polynomial $f(x_1, x_2, \ldots, x_r)$ and wish to know if there exist a $(u_1, u_2, \ldots, u_s)$ and $h_1(x_1, x_2, \ldots, x_r), \ldots, h_s(x_1, x_2, \ldots, x_r)$ such that

$$f(x_1, x_2, \ldots, x_r) = g(h_1(x_1, x_2, \ldots, x_r), \ldots, h_s(x_1, x_2, x_r).$$

hough even partial solutions would be an aid to algebraic simplification and evaluation roblems, we know of no results in this direction.

Others have attacked similar problems. Trager & Yun, (1976) presented an algorithm or completing nth powers of polynomials. That is, given a polynomial $f(x)$, they etermine if it can be written in the form $a(g(x))^n + b$ for a and b elements of k. An earlier ersion of the first decomposition algorithm given here was presented by Barton & Zippel 1976).

Finally, we note some potential applications of our algorithm. In addition to solving olynomial equations it may be used for pre-conditioning polynomials which are to be peatedly evaluated, e.g. in coding theory and numerical analysis. Also, the ecomposition of a polynomial gives the complete lattice structure of the subfields of the enus 0 curve $y = f(x)$.

The presentation in this paper was substantially improved by comments from Gerry Royalance. upport for the preparation of this paper was provided under DARPA contract 00014-80-C-0622.

References

rton, D. R., Zippel, R. E. (1976). Polynomial decomposition. *Proceedings of SYMSAC 76*, pp. 356–358.
orey, F., Whaples, G. (1974). Prime and composite polynomials. *J. Algebra* **28**, 88–101.
gström, H. T. (1941). Polynomial substitutions. *Amer. J. Math.* **63**, 249–255.
ied, M. D., MacRae, R. E. (1969*a*). On the invariance of chains of fields. *Ill. J. Math.* **13**.
ied, M. D., MacRae, R. E. (1969*b*). On curves with separate variables. *Math. Ann.* **180**, 220–226.
ied, M. D. (1974). Arithmetical properties of function fields (II). The generalized Schur problem. *Acta Arith.* **XXV**, 225–258.
usch, H., Nöbauer, W. (1973). *Algebra of Polynomials*. Amsterdam: North-Holland Pub. Co.
evi, H. (1942). Composite polynomials with coefficients in an arbitrary field of characteristic zero. *Amer. J. Math.* **64**, 389–400.
ll, J. F. (1926). Prime and composite polynomials. *Trans. Amer. Math. Soc.* **23**, 51–66.
hinzel, A. (1982). *Selected Topics on Polynomials*. Ann Arbor: University of Michigan Press.
rager, B. M., Yun, D. Y. Y. (1976). Completing nth powers of polynomials. *Proceedings of SYMSAC 76*, pp. 351–355.

J. Symb. Comp. 1 (1985) 169.

Decreasing the Nesting Depth of Expressions Involving Square Roots

ALLAN BORODIN,† RONALD FAGIN,‡ JOHN E. HOPCROFT§ AND MARTIN TOMPA ‖

† *Department of Computer Science, University of Toronto, Toronto, Ontario, Canada M5S 1A4;*
‡ *IBM Research Laboratory, 5600 Cottle Road, San Jose, California 95193, U.S.A.;*
§ *Department of Computer Science, Cornell University, Ithaca, New York 14853, U.S.A.; and*
‖ *Department of Computer Science, University of Washington, Seattle, Washington 98195, U.S.A.*

We develop the theory for decreasing the depth of nesting in expressions that contain square roots. We show exactly when fourth roots enable denesting of expressions not denestable with square roots only. When we restrict our attention to denesting over the real numbers, we show in fact that no roots other than square roots or fourth roots are ever useful for denesting expressions containing square roots only, thus characterising the denestable expressions in this case.

We then proceed to describe new algorithms that accomplish such denesting.

Algebraic Simplification and Denesting

Algebraic simplification is an important component of any symbolic manipulation system. In a general setting the simplification problem is undecidable (Richardson, 1968). For important special cases, however, simplification may be feasible. Here we are concerned with simplification of expressions containing nested radicals; in particular, we wish to decrease the depth of nesting whenever possible. This process is called denesting. Two simple examples of formulas that denest are

$$\sqrt{5+2\sqrt{6}} = \sqrt{2}+\sqrt{3}$$

and

$$\sqrt{3+2\sqrt{3}} = \tfrac{1}{4}\sqrt[4]{12}(2+\sqrt{12}).$$

Earlier work was done by Caviness & Fateman (1976) and Zippel (1977). Zippel showed how to denest certain expressions containing square roots using only square roots in the denesting, but left open the question of whether arbitrary roots are of any use in denesting such expressions. In section 1 we develop the theory for denesting expressions that contain square roots. We show exactly when fourth roots enable denesting of expressions not denestable with square roots only. When we restrict our attention to denesting over the real numbers, we show in fact that no roots other than square roots or fourth roots are ever useful for denesting expressions containing square roots only, thus characterising the denestable expressions in this important case.

This research was funded in part by the Natural Sciences and Engineering Research Council of Canada under operating grant A-7631, and by the National Science Foundation under grants MCS-8101220, MCS-8110089, and MCS-8301212.

0747-7171/85/020169+20 $03.00/0

In sections 2–6 we then proceed to describe new algorithms that denest these expressions. The algorithm of section 2 handles certain simple expressions of arbitrary nesting depth, such as the following example of depth 3

$$\sqrt{16-2\sqrt{29}+2\sqrt{55-10\sqrt{29}}} = \sqrt{5}+\sqrt{11-2\sqrt{29}}.$$

Section 4 adds the capability of fourth roots to the denested formulas. Section 5 extends the algorithm to handle rational combinations of such nested expressions, for instance

$$\sqrt{1+\sqrt{3}+\sqrt{3+3\sqrt{3}}-\sqrt{10+6\sqrt{3}}},$$

which denests to 0, despite the fact that no individual term denests. Finally, section 6 presents an algorithm that denests more complicated doubly nested formulas, such as

$$\sqrt{(112+70\sqrt{2})+(46+34\sqrt{2})\sqrt{5}} = (5+4\sqrt{2})+(3+\sqrt{2})\sqrt{5}.$$

If K is a field, then $K(a_1, \ldots, a_n)$ denotes the smallest field containing K and $a_1, \ldots, a_n$, and is called an *extension* of K. If $K(a_1, \ldots, a_n) \subseteq R$, where R is the field of real numbers, then the extension is called a *real* extension. The *degree* of $K(a_1, \ldots, a_n)$ over K, denoted $[K(a_1, \ldots, a_n) : K]$, is the dimension of $K(a_1, \ldots, a_n)$ as a vector space over K. The reader is referred to Herstein (1975) for basic facts concerning extension fields.

A *formula* over a field K and its *depth* of nesting over K are defined as follows:

1. an element of K is a formula of depth 0 over K,
2. an arithmetic combination of formulas A and B is a formula whose depth over K is $\max(\text{depth}(A), \text{depth}(B))$, and
3. a root of a formula A is a formula whose depth over K is $1+\text{depth}(A)$.

Throughout the paper the word root always refers to the principal (positive real) root. A formula can be *denested* over K if it is equal to another formula of lesser depth over K. If K is not explicitly mentioned, the words depth and denest refer respectively to depth over Q, and denest over Q, where Q is the field of rational numbers.

1. Basic Theorems for Denesting

In this section we ask when a formula of depth 2 over some field K can be rewritten as a formula of depth 1 over K. Theorem 1, most of which is implicit in Zippel (1977), shows that if $\sqrt{a+b\sqrt{r}}$ can be denested using only square roots, then it can be denested introducing only one new square root. In addition, Theorem 1 gives a necessary and sufficient condition describing when such a formula can be denested.

Theorem 2 is analogous to Theorem 1 and covers the case when fourth roots help in denesting. Theorem 3 shows that roots other than square roots or fourth roots never help if all roots involved are real.

THEOREM 1. *Let $Q \subseteq K$ and let $a, b, r \in K$ with $\sqrt{r} \notin K$. Then the following three statements are equivalent.*

(1) $\sqrt{a+b\sqrt{r}} \in K(\sqrt{r}, \sqrt{a_1}, \ldots, \sqrt{a_k})$, for some $a_1, \ldots, a_k \in K$.

(2) $\sqrt{s(a+b\sqrt{r})} \in K(\sqrt{r})$, for some $s \in K-\{0\}$.

(3) $\sqrt{a^2-b^2r} \in K$.

PROOF. (1)⇒(2): This part is by induction on k.

Basis ($k=0$): Choose $s=1$.

Induction ($k>0$): Without loss of generality assume

$$\sqrt{a_k} \notin K(\sqrt{r}, \sqrt{a_1}, \ldots, \sqrt{a_{k-1}}).$$

Let

$$\sqrt{a+b\sqrt{r}} = c+d\sqrt{a_k},$$

where c and d are in $K(\sqrt{r}, \sqrt{a_1}, \ldots, \sqrt{a_{k-1}})$. Then

$$a+b\sqrt{r} = (c^2+d^2a_k)+2cd\sqrt{a_k}.$$

Since $a+b\sqrt{r}$, $c^2+d^2a_k$ and $2cd$ are each in $K(\sqrt{r}, \sqrt{a_1}, \ldots, \sqrt{a_{k-1}})$ but $\sqrt{a_k}$ is not, $2cd=0$.

Case 1 ($c=0$): Then $\sqrt{a_k(a+b\sqrt{r})} = a_k d$ is an element of $K(\sqrt{r}, \sqrt{a_1}, \ldots, \sqrt{a_{k-1}})$. By the induction hypothesis there is an s in $K-\{0\}$ such that $\sqrt{sa_k(a+b\sqrt{r})}$ is an element of $K(\sqrt{r})$. Let $\hat{s} = sa_k$. Then $\sqrt{\hat{s}(a+b\sqrt{r})}$ is in $K(\sqrt{r})$ and $\hat{s}$ is in $K-\{0\}$.

Case 2 ($d=0$): Then $\sqrt{a+b\sqrt{r}} = c \in K(\sqrt{r}, \sqrt{a_1}, \ldots, \sqrt{a_{k-1}})$. By the induction hypothesis there is an s in $K-\{0\}$ such that $\sqrt{s(a+b\sqrt{r})}$ is in $K(\sqrt{r})$.

(2)⇒(3):

Let $s(a+b\sqrt{r}) = (c+d\sqrt{r})^2 = (c^2+d^2r)+2cd\sqrt{r}$, where $c, d \in K$. Since $\sqrt{r}$ is not in K, $sa = c^2+d^2r$ and $sb = 2cd$. If $d=0$, then $b=0$ and so $\sqrt{a^2-b^2r} = a \in K$. Otherwise, eliminating c,

$$sa = \frac{s^2b^2}{4d^2} + d^2r,$$

or

$$b^2s^2 - 4ad^2s + 4d^4r = 0.$$

Since $s \in K$, the discriminant $16a^2d^4 - 16b^2d^4r$ must be a perfect square in K, or $\sqrt{a^2-b^2r} \in K$.

(3)⇒(1):

Let $d = \sqrt{a^2-b^2r} \in K$ and $s = 2(a+d)$. If $s=0$, then $a = -d = -\sqrt{a^2-b^2r}$, so $a^2 = a^2-b^2r$, from which $b=0$ follows, and (1) holds. Therefore, assume $s \neq 0$. Then

$$\sqrt{a+b\sqrt{r}} = \frac{s+2b\sqrt{r}}{2\sqrt{s}} \in K(\sqrt{r}, \sqrt{s}),$$

since

$$\left(\frac{s+2b\sqrt{r}}{2\sqrt{s}}\right)^2 = \frac{s^2+4b^2r+4bs\sqrt{r}}{4s} = \frac{4(a+d)^2+4(a^2-d^2)+8b(a+d)\sqrt{r}}{8(a+d)}$$

$$= \tfrac{1}{2}[(a+d)+(a-d)+2b\sqrt{r}] = a+b\sqrt{r}. \quad \square$$

It is possible that a doubly nested expression $\sqrt{a+b\sqrt{r}}$ may not denest using only square roots but may denest if the fourth root of r is introduced. An example is

$$\sqrt{4+3\sqrt{2}} = \sqrt[4]{2}(1+\sqrt{2}).$$

Applying the test "$\sqrt{a^2-b^2r} \in Q$" of Theorem 1 to $\sqrt{4+3\sqrt{2}}$ shows that it cannot be denested using square roots only. Theorem 2 gives the conditions under which the fourth root of r helps.

THEOREM 2. *Let $Q \subseteq K$ and let $a, b, r \in K$ with $\sqrt{r} \notin K$. Then the following three statements are equivalent.*

(1) $\sqrt{a+b\sqrt{r}} \in K(\sqrt[4]{r}, \sqrt{a_1}, \ldots, \sqrt{a_k})$, for some $a_1, \ldots, a_k \in K$.

(2) $\sqrt[4]{r}\sqrt{s}\sqrt{a+b\sqrt{r}} \in K(\sqrt{r})$ or $\sqrt{s}\sqrt{a+b\sqrt{r}} \in K(\sqrt{r})$, for some $s \in K-\{0\}$.

(3) $\sqrt{r(b^2r-a^2)} \in K$ or $\sqrt{a^2-b^2r} \in K$.

PROOF. (1) $\Rightarrow$ (2):

Assume $\sqrt{a+b\sqrt{r}}$ is not in $K(\sqrt{r}, \sqrt{a_1}, \ldots, \sqrt{a_k})$, since otherwise $\sqrt{s}\sqrt{a+b\sqrt{r}} \in K(\sqrt{r})$ follows from Theorem 1. Let $a+b\sqrt{r} = (c^2+d^2\sqrt{r})+2cd\sqrt[4]{r}$, where c and d are in $K(\sqrt{r}, \sqrt{a_1}, \ldots, \sqrt{a_k})$. Furthermore, we can assume $d \neq 0$ and $\sqrt[4]{r} \notin K(\sqrt{r}, \sqrt{a_1}, \ldots, \sqrt{a_k})$, since $\sqrt{a+b\sqrt{r}}$ is in $K(\sqrt[4]{r}, \sqrt{a_1}, \ldots, \sqrt{a_k}) - K(\sqrt{r}, \sqrt{a_1}, \ldots, \sqrt{a_k})$. Hence $c=0$ and $\sqrt{a+b\sqrt{r}} = d\sqrt[4]{r}$. Multiplication by $\sqrt[4]{r}$ yields

$$d\sqrt{r} = \sqrt[4]{r}\sqrt{a+b\sqrt{r}} = \sqrt{br+a\sqrt{r}}.$$

This means $\sqrt{br+a\sqrt{r}}$ is in $K(\sqrt{r}, \sqrt{a_1}, \ldots, \sqrt{a_k})$. By Theorem 1, $\sqrt{s(br+a\sqrt{r})}$ is in $K(\sqrt{r})$ for some s in $K-\{0\}$. That is $\sqrt[4]{r}\sqrt{s}\sqrt{a+b\sqrt{r}}$ is in $K(\sqrt{r})$.

(2) $\Rightarrow$ (3):

If $\sqrt{s}\sqrt{a+b\sqrt{r}}$ is in $K(\sqrt{r})$, then $\sqrt{a^2-b^2r}$ is in K by Theorem 1. Otherwise $\sqrt[4]{r}\sqrt{s}\sqrt{a+b\sqrt{r}}$ is in $K(\sqrt{r})$, or equivalently $\sqrt{s(br+a\sqrt{r})}$ is in $K(\sqrt{r})$. By Theorem 1, $\sqrt{b^2r^2-a^2r}$ is in K.

(3) $\Rightarrow$ (1):

If $\sqrt{a^2-b^2r}$ is in K, then $\sqrt{a+b\sqrt{r}}$ is in $K(\sqrt{r}, \sqrt{s})$, by Theorem 1. By Theorem 1, $\sqrt{b^2r^2-a^2r}$ in K implies $\sqrt{br+a\sqrt{r}} = \sqrt[4]{r}\sqrt{a+b\sqrt{r}}$ is in $K(\sqrt{r}, \sqrt{s})$ for some s in K, so $\sqrt{a+b\sqrt{r}}$ is in $K(\sqrt[4]{r}, \sqrt{s})$. □

Theorems 1 and 2 will be used in algorithms that denest formulas. Theorem 2 raises the possibility that roots other than square roots may be useful. In Theorem 3 we show that, with the exception of the fourth root of r, no additional roots are useful and that Theorems 1 and 2 cover all possible denestings.

Whereas Theorems 1 and 2 hold for arbitrary extensions, Theorem 3 holds only for real extensions of K. For example, $\sqrt{1+\sqrt{-1}}$ does not denest using only square roots and fourth roots, but can be denested using an eighth root: $\sqrt{1+\sqrt{-1}} = \sqrt[8]{-4}$. One may question whether this is really a simplification, or whether one would prefer to treat $\sqrt[8]{-4}$ as $\sqrt{\sqrt{\sqrt{-4}}}$ and denest to $\sqrt{1+\sqrt{-1}}$.

To simplify the proof of Theorem 3 we first prove five lemmas. Given a real extension of K in a particular normal form, Lemma 1 states that for any x in K with $\sqrt{x}$ in the extension, $\sqrt{x}$ can be expressed in a simple form. Again, given the normal form, Lemma 2 states that certain roots of elements of K are not in the extension. Lemmas 3 and 4 establish that certain real extensions of K can be placed in the normal form. Lemma 5 deals with the degree of an extension obtained by adding a prime root.

For similar results in the case when the underlying field K is Q, the reader is referred to Besicovitch (1940) and Richards (1974).

LEMMA 1. *Let K be a real extension of Q, and let $a_1, \ldots, a_{n+1}$ be positive elements of K. Let $r_1, \ldots, r_n \geq 2$ be powers of 2, where $r_i \neq 1$ for $1 \leq i \leq n$. Suppose that*

$$[K(\sqrt[r_1]{a_1}, \ldots, \sqrt[r_n]{a_n}) : K] = \prod_{i=1}^{n} r_i$$

and that the r_i are minimal powers of 2 such that $\sqrt{a_{n+1}}$ is in $K(\sqrt[r_1]{a_1}, \ldots, \sqrt[r_n]{a_n})$. Then

$$r_1 = r_2 = \cdots = r_n = 2 \quad \text{and} \quad \sqrt{a_{n+1}} = a\sqrt{a_1 \cdots a_n},$$

for some a in K.

PROOF. The lemma is proved by induction on n.

Basis ($n=0$): Let $a = \sqrt{a_{n+1}}$.

Induction ($n>0$): Let $L = K(a_1^{1/r_1}, \ldots, a_{n-1}^{1/r_{n-1}}, a_n^{2/r_n})$ and write $\sqrt{a_{n+1}} = ba_n^{1/r_n} + c$, with b and c in L. By the minimality of r_n, $b \neq 0$. By squaring,

$$a_{n+1} = b^2a_n^{2/r_n} + c^2 + 2bca_n^{1/r_n}.$$

Since $b \neq 0$ and a_n^{1/r_n} is not in L, we conclude that $c=0$. Assume now that $r_n = 2$. Then $\sqrt{a_{n+1}a_n} = ba_n$, so

$$\sqrt{a_{n+1}a_n} \in K(a_1^{1/r_1}, \ldots, a_{n-1}^{1/r_{n-1}}).$$

Therefore, by the induction hypothesis, each $r_i = 2$, and there is $a \in K$ such that

$$\sqrt{a_{n+1}a_n} = a\sqrt{a_1 \cdots a_{n-1}}.$$

Hence

$$\sqrt{a_{n+1}} = aa_n^{-1}\sqrt{a_1 \cdots a_n}.$$

Thus assume $r_n \geq 4$. Let $\hat{L} = K(a_1^{1/r_1}, \ldots, a_{n-1}^{1/r_{n-1}}, a_n^{4/r_n})$ and write

$$\sqrt{a_{n+1}} = ba_n^{1/r_n} = (d + ea_n^{2/r_n})a_n^{1/r_n},$$

where $d, e \in \hat{L}$. Again, by squaring

$$a_{n+1} = (d^2 + e^2a_n^{4/r_n})a_n^{2/r_n} + 2dea_n^{4/r_n}.$$

Since $[L : \hat{L}] = 2$, a_n^{2/r_n} is not in $\hat{L}$, and therefore $d^2 + e^2a_n^{4/r_n} = 0$. This contradicts the hypothesis that a_n is positive, so the assumption $r_n \geq 4$ is false. □

Observe that the hypothesis that K is a real extension is needed, since

$$\sqrt{2} = (1-\sqrt{-1})\sqrt[4]{-1} \in Q(\sqrt[4]{-1}) - Q(\sqrt{-1}).$$

LEMMA 2. *Let K be a real extension of Q and let a be a positive element of K. Let $r \geq 2$ be a power of 2. Suppose $[K(a^{1/r}) : K] = r$. Then $a^{1/2r} \notin K(a^{1/r})$.*

PROOF. Suppose that $a^{1/2r}$ is in $K(a^{1/r})$. Write $a^{1/2r} = ba^{1/r} + c$, where b and c are in $K(a^{2/r})$.

Squaring we get

$$a^{1/r} = b^2 a^{2/r} + c^2 + 2bca^{1/r}.$$

Since $a^{1/r}$ is not in $K(a^{2/r})$ we conclude that $2bc = 1$, so $b^2a^{2/r} + c^2 = 0$. Thus either $b = c = 0$ or $a^{2/r} = -(c/b)^2$, in either case a contradiction. □

LEMMA 3. *Let K be a real extension of Q. Let $n_1, \ldots, n_k$ be powers of 2 and let*

$$n = \prod_{i=1}^{k} n_i.$$

Let $a_1, \ldots, a_k$ be positive elements of K. Let

$$L = K(\sqrt[n_1]{a_1}, \ldots, \sqrt[n_k]{a_k}).$$

If $[L:K] \neq n$, then we can find positive $b_1, \ldots, b_k$ in K, and integers $m_1, \ldots, m_k$, each a power of 2, such that

$$L = K(\sqrt[m_1]{b_1}, \ldots, \sqrt[m_k]{b_k}) \quad \text{and} \quad m = \prod_{i=1}^{k} m_i < n.$$

PROOF. Renumber so that $n_1 \geq n_2 \geq \ldots \geq n_k$. Consider the sequence of extensions

$$K,\ K(\sqrt{a_1}),\ K(\sqrt[4]{a_1}),\ \ldots,\ K(\sqrt[n_1]{a_1}),\ K(\sqrt[n_1]{a_1}, \sqrt{a_2}),\ K(\sqrt[n_1]{a_1}, \sqrt[4]{a_2}),\ \ldots.$$

Consider the first step at which the extension is not proper. Then by Lemma 2 we must be adding $\sqrt{a_i}$ for some i and

$$\sqrt{a_i} \in K(a_1^{1/n_1}, \ldots, a_{i-1}^{1/n_{i-1}}).$$

By Lemma 1, $\sqrt{a_i} = b\sqrt{a_{i_1} \cdots a_{i_r}}$ for some $b \in K$ and $i_1, \ldots, i_r \in \{1, 2, \ldots, i-1\}$. Let

$$\hat{L} = K(a_1^{1/n_1}, \ldots, a_{i-1}^{1/n_{i-1}}, b^{2/n_i}, a_{i+1}^{1/n_{i+1}}, \ldots, a_k^{1/n_k}).$$

Since

$$b^{2/n_i} = \left(\frac{a_i}{a_{i_1} \cdots a_{i_r}}\right)^{1/n_i},$$

it is clear that $L = \hat{L}$ and $m = n/2$. □

LEMMA 4. *Given L as in Lemma 3 we can find positive $b_1, \ldots, b_k$ in K and integers $m_1, \ldots, m_k$, each a power of 2, such that*

$$L = K(\sqrt[m_1]{b_1}, \ldots, \sqrt[m_k]{b_k}) \quad \text{and} \quad [L:K] = \prod_{i=1}^{k} m_i.$$

PROOF. This follows from repeated application of Lemma 3. □

LEMMA 5. (*Abel (Schinzel, 1982, p. 91)*). *Let L be a real extension of the rationals, p an odd prime, b a real number, and b^p an element of L. Then, provided $b \notin L$, $[L(b):L] = p$.*

PROOF. Any irreducible factor of $x^p - b^p$ over the reals is $x - b$ or of the form $x^2 - (\omega + \bar{\omega})bx + b^2$, where ω and $\bar{\omega}$ are complex conjugates. (This is because the complex roots of $x^p - b^p = 0$ are of the form ωb, where ω is a pth root of 1, and because

$$(x - \omega b)(x - \bar{\omega} b) = x^2 - (\omega + \bar{\omega})b + b^2).$$

Suppose $x^p - b^p$ could be factored over L. Since L is a real extension, any proper factor of $x^p - b^p$ over L must be a product of some proper subset of its factors over the reals, and hence have a constant term of the form b^i, $1 \leq i < p$. Since p is prime, $b^i \in L$ implies $b \in L$, a contradiction. Thus we conclude that $x^p - b^p$ is irreducible over L and hence $[L(b):L] = p$ (see Herstein, 1975). □

As an example of the importance of the real assumptions, let

$$b = \frac{1 + \sqrt{-3}}{2},$$

and $p = 3$. Then $b^p = -1$, but $Q(b) = Q(\sqrt{-3})$, so $[Q(b):Q] = 2$.

THEOREM 3. *Let K be a real extension of Q and let $a, b, r \in K$ with $\sqrt{r} \notin K$. Let $n_1, \ldots, n_k \geq 1$ and $a_1, \ldots, a_k \in K$ be positive. If*

$$\sqrt{a + b\sqrt{r}} \in K(\sqrt[n_1]{a_1}, \ldots, \sqrt[n_k]{a_k})$$

then

$$\sqrt{a + b\sqrt{r}} \in K(\sqrt[4]{r}, \sqrt{p_1}, \ldots, \sqrt{p_k})$$

for some positive $p_1, \ldots, p_k$ in K.

PROOF. Let $n_1, \ldots, n_k$ be minimal such that $\sqrt{a + b\sqrt{r}}$ is in $K(a_1^{1/n_1}, \ldots, a_k^{1/n_k})$. Suppose n_k were divisible by an odd prime p. Let $L = K(a_1^{1/n_1}, \ldots, a_{k-1}^{1/n_{k-1}})$. By the minimality of n_k,

$$L(a_k^{p/n_k}) \subsetneq L(a_k^{p/n_k}, \sqrt{a + b\sqrt{r}}) \subseteq L(a_k^{1/n_k}).$$

But $[L(a_k^{p/n_k}, \sqrt{a + b\sqrt{r}}) : L(a_k^{p/n_k})]$ is 2 or 4 and $[L(a_k^{1/n_k}) : L(a_k^{p/n_k})] = p$ (by Lemma 5), a contradiction (see Herstein, 1975). Thus each n_i is a power of 2.

Now let $L = K(a_1^{1/n_1}, \ldots, a_k^{1/n_k})$. By Lemma 4 there exist positive $b_1, \ldots, b_k$ in K and integers $m_1, \ldots, m_k$, each a power of 2 and each at least 2, such that $L = K(b_1^{1/m_1}, \ldots, b_k^{1/m_k})$ and

$$[L:K] = \prod_{i=1}^{k} m_i.$$

Assume $m_1, \ldots, m_k$ are the least powers of 2 such that $\sqrt{a + b\sqrt{r}}$ is still in $K(b_1^{1/m_1}, \ldots, b_k^{1/m_k})$; if not, decrease them accordingly and redefine L. Since $\sqrt{r} \in L$, Lemma 1 states that $\sqrt{r} = c\sqrt{b_1 \cdots b_j}$ for some $c \in K$. Assume that $m_1 \leq m_2 \leq \ldots \leq m_j$, possibly renumbering the b_i. Let $r_1 = b_1 \cdots b_j$. Now $L = K(r_1^{1/m_1}, b_2^{1/m_2}, \ldots, b_k^{1/m_k})$. Suppose by way of contradiction that $m_1 \geq 8$, or $i > 1$ and $m_i \geq 4$, and denote r_1^{1/m_1} (in the first case) or b_i^{1/m_i} (in the second) by $p^{1/m}$. Let $\hat{L} = K(b_2^{1/m_2}, \ldots, b_k^{1/m_k})$ (in the first case) and

$$\hat{L} = K(r_1^{1/m_1}, \ldots, b_{i-1}^{1/m_{i-1}}, b_{i+1}^{1/m_{i+1}}, \ldots, b_k^{1/m_k})$$

(in the second). Observe that $\sqrt{r} = c\sqrt{r_1}$ is in $\hat{L}(p^{4/m})$.

By the minimality of m, $\sqrt{a + b\sqrt{r}}$ is in $\hat{L}(p^{1/m}) - \hat{L}(p^{2/m})$. That is,

$$\sqrt{a + b\sqrt{r}} = c + dp^{1/m},$$

where c and d are in $\hat{L}(p^{2/m})$ and $d \neq 0$. By squaring,

$$a + b\sqrt{r} = c^2 + d^2p^{2/m} + 2cdp^{1/m}.$$

Since all but the last term $2cdp^{1/m}$ are in $\hat{L}(p^{2/m})$, it follows that $cd=0$, so $c=0$. Write

$$\sqrt{a+b\sqrt{r}} = dp^{1/m} = (e+fp^{2/m})p^{1/m},$$

where e and f are in $\hat{L}(p^{4/m})$. Again, by squaring,

$$a+b\sqrt{r} = (e^2+f^2p^{4/m})p^{2/m}+2efp^{4/m}.$$

All subformulas of this equation are in $\hat{L}(p^{4/m})$ except for $p^{2/m}$, since

$$[L:K] = \prod_{i=1}^{h} m_i.$$

Thus $e^2+f^2p^{4/m}=0$ implying that $e=f=0$ or $p^{4/m}=-e^2/f^2$, a contradiction. Thus, m_1 is in $\{2,4\}$ and $m_i=2$ for $i>1$, i.e.

$$\sqrt{a+b\sqrt{r}} \in K(\sqrt[4]{r_1}, \sqrt{b_2}, \ldots, \sqrt{b_k}).$$

Since

$$\sqrt[4]{r} = \sqrt{c}\,\sqrt[4]{r_1}, \sqrt{a+b\sqrt{r}} \in K(\sqrt[4]{r}, \sqrt{c}, \sqrt{b_2}, \ldots, \sqrt{b_k}). \quad \square$$

2. Denesting Using Square Roots Only

The goal of the remaining sections is to present new algorithms that, whenever possible, rewrite a given formula with positive radicands as one that has a lower nesting depth over Q. (The algorithms will sometimes, but not always, denest formulas that have negative radicands.) Consider the following tower of fields:

$$K_0 = Q \quad \text{and, for all } i \geqslant 0,\ K_{i+1} = K_i(\sqrt{r_i}),$$

where $r_i \in K_i$. Every element of K_i has depth at most i over Q. Consider $\sqrt{r_n}$, where $r_n \in K_n$. This section presents an algorithm, based on Theorem 1, that rewrites $\sqrt{r_n}$ as a formula of depth at most n over Q involving only square roots, whenever such a formula exists (even if the formulas have negative radicands).

Although the fundamental ideas appear in this algorithm, it is incomplete in three respects. Firstly, it does not permit the fourth roots of Theorem 2; this is rectified in Section 4. Secondly, the algorithm denests only single radicals from K_{n+1} rather than arbitrary elements of K_{n+1}; this is rectified in section 5. Finally, not every formula of depth n over Q can be expressed in the form of elements of K_n; in particular, elements of K_n can be written using only n distinct radicals. Section 6 rectifies this only partially, by describing how to denest arbitrary formulas of depth 2 over Q, and explaining why this technique does not generalise to depths greater than 2.

Let $r_n = a+b\sqrt{r_{n-1}}$, where $a, b, r_{n-1} \in K_{n-1}$. There are two ways in which $\sqrt{r_n}$ might denest using square roots only. The first way is that $\sqrt{r_{n-1}}$ might denest. The second possibility is to find some field K containing only elements of depth $n-1$, containing a, b, and r_{n-1}, and in which $a^2-b^2r_{n-1}$ is a perfect square. Then $\sqrt{r_n}$ denests by Theorem 1. In order to find the K in which $a^2-b^2r_{n-1}$ is a perfect square we attempt to denest $\sqrt{a^2-b^2r_{n-1}}$. If we are successful we will find a field $\hat{K}$, all of whose elements have depth at most $n-2$ and an element s in $\hat{K}$ such that $\sqrt{a^2-b^2r_{n-1}}$ is in $\hat{K}(\sqrt{s}, \sqrt{r_{n-2}})$. The desired K is then $\hat{K}(\sqrt{s}, \sqrt{r_{n-2}})$.

Since $\sqrt{r_{n-1}}$ and $\sqrt{a^2-b^2r_{n-1}}$ each have depth less than $\sqrt{r_n}$, this suggests a recursive algorithm: to denest $\sqrt{r_n}$, try to denest each of $\sqrt{r_{n-1}}$ and $\sqrt{a^2-b^2r_{n-1}}$. If either succeeds, then $\sqrt{r_n}$ can be denested. If neither succeeds, $\sqrt{r_n}$ cannot be denested using square roots alone.

This algorithm has a serious shortcoming: it requires exponentially many recursive invocations. However, closer inspection reveals that there are not 2^i distinct arguments at the ith level of recursion, but only $i+1$, as $\sqrt{r_{n-i}}$ recurs repeatedly. Continuing the previous discussion of denesting $\sqrt{r_n}$, the first level of recursion attempted to denest $\sqrt{r_{n-1}}$ and $\sqrt{a^2-b^2r_{n-1}}$. Each of these radicands is in K_{n-1}, so

$$r_{n-1} = c+d\sqrt{r_{n-2}}, \quad \text{and} \quad a^2-b^2r_{n-1} = e+f\sqrt{r_{n-2}},$$

for some $c, d, e, f, r_{n-2} \in K_{n-2}$. Thus, the only formulas that need be denested at the second level of recursion are $\sqrt{r_{n-2}}$, $\sqrt{c^2-d^2r_{n-2}}$, and $\sqrt{e^2-f^2r_{n-2}}$. In general, only one more argument will be added at each level of recursion.

This motivates the algorithm of this section. We assume for the remainder of this paper that there is an appropriate data type *formula* for representing formulas. The recursive algorithm takes as input 2 integers, n and m, and a list of m formulas, each of depth n over Q. The algorithm attempts to denest each of these, returning for each input formula either a denested version, or the input formula itself if it was unsuccessful.

Notice in the algorithm that it is assumed that each input formula shares the same value of $\sqrt{r}$ if $n \geqslant 2$; that is, the ith formula is of the form $\sqrt{a_i+b_i\sqrt{r}}$.

```
function DENEST (integer, n, m,
        array [1 . . . m] of formula nested)
    returns array [1 . . . m] of formula;
begin
    declare N, d: array [1 . . . m+1] of formula;
    declare s: formula;
    if n = 1
        then for i from 1 to m do
            if nested[i] = √x² for some positive x ∈ Q
                then DENEST[i] ← x
                else DENEST[i] ← nested[i]
        else begin
            assume nested[i] = √(a_i + b_i√r), 1 ≤ i ≤ m;
            for i from 1 to m do N[i] ← √(a_i² − b_i²r);
            N[m+1] ← √r;
            d ← DENEST(n−1, m+1, N);
            if N[m+1] ≠ d[m+1] comment: √r denested;
                then for i from 1 to m do
                    DENEST[i] ← √(a_i + b_i*d[m+1])
                else for i from 1 to m do
                    if N[i] = d[i]
                        then DENEST[i] ← nested[i]
                        else begin
                            s ← 2(a_i + d[i]);
                            DENEST[i] ← (s + 2b_i√r) / 2√s
                        end
        end
end DENEST.
```

Consider the following example due to Shanks, see Zippel (1977), the object of which is to denest

$$\sqrt{16-2\sqrt{29}+2\sqrt{55-10\sqrt{29}}}$$

Here

$$K_1 = Q(\sqrt{29})$$

and

$$K_2 = K_1(\sqrt{55-10\sqrt{29}}),$$

satisfying $16-2\sqrt{29} \in K_1$, $2 \in K_1$, and $55-10\sqrt{29} \in K_1$. The first call is

$$\text{DENEST}(3, 1, [\sqrt{16-2\sqrt{29}+2\sqrt{55-10\sqrt{29}}}]).$$

This produces the two formulas

$$\sqrt{(16-2\sqrt{29})^2-2^2(55-10\sqrt{29})} = \sqrt{152-24\sqrt{29}} \quad \text{and} \quad \sqrt{55-10\sqrt{29}},$$

resulting in the second call of

$$\text{DENEST}(2, 2, [\sqrt{152-24\sqrt{29}}, \sqrt{55-10\sqrt{29}}]).$$

This in turn produces the three formulas

$$\sqrt{152^2-24^2\times 29} = \sqrt{6400}, \quad \sqrt{55^2-10^2\times 29} = \sqrt{125}, \quad \text{and} \quad \sqrt{29},$$

resulting in the third call of DENEST$(1, 3, [\sqrt{6400}, \sqrt{125}, \sqrt{29}])$. The basis portion of the algorithm succeeds in denesting the first of these, returning the list $[80, \sqrt{125}, \sqrt{29}]$. The invoking procedure can use this to denest its first input $\sqrt{152-24\sqrt{29}}$ as follows

$$s = 2(152+80) = 464,$$

$$\sqrt{152-24\sqrt{29}} = \frac{464-2\times 24\sqrt{29}}{2\sqrt{464}} = 2\sqrt{29}-6.$$

(The last simplification would not be done by the algorithm, but it makes the example clearer. Notice that it does not affect the depth of nesting.)

Thus, this invocation returns the list $[2\sqrt{29}-6, \sqrt{55-10\sqrt{29}}]$. Finally, the initial invocation completes as follows

$$s = ((16-2\sqrt{29})+(2\sqrt{29}-6)) = 20,$$

$$\sqrt{16-2\sqrt{29}+2\sqrt{55-10\sqrt{29}}} = \frac{20+2\times 2\sqrt{55-10\sqrt{29}}}{2\sqrt{20}} = \sqrt{5}+\sqrt{11-2\sqrt{29}},$$

returning the singleton list $[\sqrt{5}+\sqrt{11-2\sqrt{29}}]$.

The correctness of this section's algorithm is the subject of the next theorem.

THEOREM 4. *Let $m \geq 1$ and $n \geq 1$. Given m formulas $\sqrt{f_1}, \sqrt{f_2}, \ldots, \sqrt{f_m}$, where $f_1, f_2, \ldots, f_m \in K_{n-1}$, procedure* DENEST *will denest $\sqrt{f_i}$ correctly, whenever $\sqrt{f_i}$ can be denested using square roots only, for all $1 \leq i \leq m$.*

PROOF. The proof is by induction on n.

Basis ($n = 1$): $f_i \in Q$, so $\sqrt{f_i}$ denests if and only if f_i is a perfect square.

Induction ($n > 1$): Let $f_i = a_i + b_i\sqrt{r_{n-2}}$, where $a_i, b_i, r_{n-2} \in K_{n-2}$, and suppose $\sqrt{f_i}$ can be denested to depth $n-1$ over Q using square roots only, that is,

$$\sqrt{a_i+b_i\sqrt{r_{n-2}}} \in K(\sqrt{c_1}, \sqrt{c_2}, \ldots, \sqrt{c_k}),$$

for some K all of whose elements are of depth at most $n-2$ over Q, such depth achieved using only square roots, and some $c_1, c_2, \ldots, c_k \in K$. We can assume that $K_{n-2} \subseteq K$. By Theorem 1, either $\sqrt{r_{n-2}} \in K$ or $\sqrt{a_i^2-b_i^2 r_{n-2}} \in K$, that is, one of these two formulas denests. Each of these two radicands is in K_{n-2}, and each of these two formulas is passed in the recursive invocation so, by the induction hypothesis, DENEST will denest one or both of them correctly. If $\sqrt{r_{n-2}}$ is denested to $r \in K$, then surely $\sqrt{a_i+b_i r}$ is a correct denesting of $\sqrt{a_i+b_i\sqrt{r_{n-2}}}$. Otherwise, suppose $\sqrt{a_i^2-b_i^2 r_{n-2}}$ is denested to $d_i \in K$. Then the correctness of the denesting

$$\sqrt{a_i+b_i\sqrt{r_{n-2}}} = \frac{s+2b^i\sqrt{r_{n-2}}}{2\sqrt{s}} \in K(\sqrt{r_{n-2}}, \sqrt{s}),$$

where $s = 2(a_i+d_i)$ is supplied in the proof of Theorem 1, part (3)⇒(1). □

What remains is a discussion of the running time of the procedure DENEST. Inspection of the procedure reveals that, as long as the formulas have size polynomial in n, the number of arithmetic operations will also be polynomial in n. Unfortunately, in general the formulas have size exponential in n, as is clear from the last two lines of the procedure: s depends on $d[i]$, the result of the recursive invocation, and there are two occurrences of s in DENEST$[i]$, the result of the current invocation. We alleviate this problem in an (unfortunately) asymmetric way: although we require the inputs to be represented as trees, we represent the outputs as acyclic directed graphs (i.e. straight-line programs). This is unfortunate in the sense that the output cannot be given as an input for another level of denesting. Notice also that, unlike the arithmetic complexity, the bit complexity is not polynomially bounded, as the number of bits doubles at each recursive call due to the squaring involved in $\sqrt{a^2-b^2 r}$.

3. Finding a Perfect Square Among Products of a Set of Numbers

Before proceeding with the next algorithm we present a straightforward technique for finding a perfect square among all possible products of a set of rational numbers. The result is used as the basis for the recursive algorithm in the next section.

Let $r_1, \ldots, r_n$ be rational numbers. Let $\hat{r}_i = r_i d_i^2$, where d_i is the denominator of r_i. (Note that $\hat{r}_i$ is the product of the numerator and denominator of r_i.) Then $\prod_{i \in S} \hat{r}_i$ is a perfect square if and only if $\prod_{i \in S} r_i$ is a perfect square. Thus, without loss of generality, let $r_1, \ldots, r_n$ be integers.

In order to find a perfect square among products of $r_1, r_2, \ldots, r_n$ the greatest common divisor of each pair is determined, resulting in a set of integers p_j that are pairwise relatively prime and for which each r_i, $1 \leq i \leq n$, can be written as a product of p_j's. Remove any p_j that is a perfect square.

Next, for $i = 1, 2, \ldots, n$, replace r_i by $p_i a_i$, where a_i is a product of p_j's, $j > i$. (This may involve renumbering the p_i's.) In order to do this, if r_i contains p_j, $j < i$, then for the least such j replace r_i by $r_i r_j$, remove any p_k^2 from this product, and iterate the process until r_i contains p_j only for $j \geq i$. (At this point renumber one such p_j to be p_i, if p_i does not

already occur in r_i.) It is straightforward to verify that such a replacement does not change the existence of a subset whose product is a perfect square, and that such a subset exists if and only if some r_i eventually takes on the value 1. The subset can be identified easily by retracing the steps.

The following application of this algorithm will be used in section 6. In the case where all $r_i > 0$ we can determine if $\sqrt{r_n}$ is in $Q(\sqrt{r_1}, \ldots, \sqrt{r_{n-1}})$ by locating perfect squares among products of $r_1, \ldots, r_n$. On encountering a perfect square, if r_n is involved, then $\sqrt{r_n}$ is in $Q(\sqrt{r_1}, \ldots, \sqrt{r_{n-1}})$. Otherwise delete some r_i from the product, and repeat the process until $\sqrt{r_n}$ is discovered to be in $Q(\sqrt{r_1}, \ldots, \sqrt{r_{n-1}})$ or there are no perfect squares remaining. In the latter case $\sqrt{r_n}$ is not in $Q(\sqrt{r_1}, \ldots, \sqrt{r_{n-1}})$ by Lemma 1.

4. Denesting Using Arbitrary Roots

Consider again the tower of fields presented in section 2 (but in this section restricted to be subfields of the reals)

$$K_0 = Q,$$

and

$$\text{for all } i \geqslant 0,\ K_{i+1} = K_i(\sqrt{r_i}),$$

where $r_i \in K_i$ is positive. The goal of this section is to generalise the algorithm of section 2 so that it rewrites its input $\sqrt{r_n}$ as a formula of depth at most n over Q, whenever such a denesting is possible, yet still has arithmetic complexity polynomial in n. By Theorem 3, the only useful additions to the algorithm of section 2 are the fourth roots of Theorem 2 (provided all subexpressions that arise are real).

Let $r_n = a + b\sqrt{r_{n-1}}$, where $a, b, r_{n-1} \in K_{n-1}$. By Theorems 1–3, $\sqrt{r_n}$ denests if and only if

(i) $\sqrt{r_{n-1}}$ denests, or
(ii) $\sqrt{a^2 - b^2 r_{n-1}}$ denests, or
(((i) $\sqrt{-r_{n-1}(a^2 - b^2 r_{n-1})}$ denests.

Unfortunately, the obvious generalisation of the algorithm of section 2 has exponential complexity, since each formula $\sqrt{a + b\sqrt{r_{n-1}}}$ in the input list gives rise to two formulas, $\sqrt{a^2 - b^2 r_{n-1}}$ and $\sqrt{-r_{n-1}(a^2 - b^2 r_{n-1})}$ in the recursive input list, in addition to the one extra formula $\sqrt{r_{n-1}}$.

The solution is to deal with the formulas of the form $\sqrt{-r_{n-1}(a^2 - b^2 r_{n-1})}$ implicitly, rather than generating them explicitly, by observing that the radicand can be expressed as the product (with a possible sign change) of the radicands of $\sqrt{r_{n-1}}$ and $\sqrt{a^2 - b^2 r_{n-1}}$. Thus, the algorithm of this section will invoke itself with exactly the same lists of formulas as the one of section 2, but rather than trying to denest each individually, it will determine whether there is a subset of the formulas whose product (with a possible sign change) denests. The justification for looking at all subsets is given below in Lemma 6, which states that a product

$$\prod_{i=1}^{k} \sqrt{a_i + b_i\sqrt{r}}$$

of input formulas denests if and only if $\sqrt{r}$ denests, or the corresponding product

$$\prod_{i=1}^{k} \sqrt{a_i^2 - b_i^2 r}$$

of formulas denests, or

$$\sqrt{-r} \prod_{i=1}^{k} \sqrt{a_i^2 - b_i^2 r}$$

denests.

In the basis, when all formulas have depth 1 over Q, the algorithm applies the procedure of section 3 to determine whether any subset of its input radicands has a product (with a possible sign change) that is a perfect square. If it finds such a subset it will identify it in the global array **factor** by setting

$$\textbf{factor}[i] = \begin{cases} 1, & \text{if the } i\text{th formula is in the subset} \\ 0, & \text{otherwise} \end{cases}$$

The pending invocations will then use Theorems 1 and 2 and Lemma 6 to denest the product specified in the array **factor**.

```
function DENEST (integer n, m,
      array [1 . . . m] of formula nested)
  returns formula;
begin
  declare N: array [1 . . . m+1] of formula;
  declare d, a, b, s: formula;
  if n = 1
      then Use the procedure of section 3 to find a perfect square x² among products of the
        m input radicands. Before entering the radicand of nested[m] into any product
        with at least one other radicand, it should be multiplied by −1. Record the
        formulas entering into the product by placing 1's into the appropriate cells of the
        array factor. Return the formula x. If no such perfect square is found, set factor
        [i] = 0 for all i and return the formula nested[m].
      else begin
        assume nested[i] = √(a_i + b_i √r), 1 ⩽ i ⩽ m;
          for i from 1 to m do N[i] ← √(a_i² − b_i² r);
          N[m+1] ← √r;
          d ← DENEST(n−1, m+1, N);
          if factor[i] = 0 for all 1 ⩽ i ⩽ m        comment: only √r could have denested;
            then DENEST ← √(a_m + b_m d)
            else begin
              define a, b by √(a + b√r) = ∏_{i=1}^{m} factor[i]* nested[i];
              if factor[m] = 1 and factor[i] = 1 for some 1 ⩽ i < m
                then begin
                  a ← −a;
                  b ← −b
                end;
              comment: √(a + b√r) denests;
              if factor [m+1] = 0
                then begin
                    comment: apply Theorem 1;
                    s ← 2(a + d);
                    DENEST ← (s + 2b√r) / (2√s);
                end
                else begin
                    comment: apply Theorem 2;
                    s ← 2(br + d);
```

DENEST $\leftarrow \dfrac{s+2a\sqrt{r}}{2\sqrt{s\sqrt[4]{r}}}$;

end

end

end

end DENEST.

Consider denesting the formula $\sqrt{3+\sqrt{5+2\sqrt{7}}}$. The depth of this formula over Q is 3, so the global declaration needed is

declare factor: *array* [1.3] *of* {0, 1},

and the first invocation is DENEST(3, 1, $[\sqrt{3+\sqrt{5+2\sqrt{7}}}]$). This gives rise to the two formulas

$$\sqrt{3^2-(5+2\sqrt{7})}=\sqrt{4-2\sqrt{7}} \quad \text{and} \quad \sqrt{5+2\sqrt{7}},$$

resulting in the second invocation DENEST(2, 2, $[\sqrt{4-2\sqrt{7}}, \sqrt{5+2\sqrt{7}}]$). This in turn produces the three formulas

$$\sqrt{4^2-2^2\times 7}=\sqrt{-12}, \quad \sqrt{5^2-2^2\times 7}=\sqrt{-3}, \quad \text{and} \quad \sqrt{7},$$

resulting in the third invocation DENEST(1, 3, $[\sqrt{-12}, \sqrt{-3}, \sqrt{7}]$). The algorithm of section 3 determines that $(-12)(-3)$ is a perfect square, sets factor to [1, 1, 0], and returns 6. In the second invocation, $a=8$ and $b=2$, since

$$\sqrt{(-1)(4-2\sqrt{7})(5+2\sqrt{7})}=\sqrt{8+2\sqrt{7}}.$$

Since factor[3] = 0 (i.e. $\sqrt{-r}$ was not involved in the product), $s=2(8+6)=28$ and

$$\sqrt{8+2\sqrt{7}}=\frac{28+4\sqrt{7}}{2\sqrt{28}}=1+\sqrt{7},$$

so $1+\sqrt{7}$ is returned to the first invocation. In this invocation, $a=3$ and $b=1$. Since factor[2] $\neq 0$ (i.e. $\sqrt{-r}$ was involved in the product), a fourth root is needed as follows

$$s=2((5+\sqrt{7})+(1+\sqrt{7}))=12+6\sqrt{7},$$

$$\sqrt{3+\sqrt{5+2\sqrt{7}}}=((12+6\sqrt{7})+6\sqrt{5+2\sqrt{7}})/(2\sqrt{12+6\sqrt{7}}\sqrt[4]{5+2\sqrt{7}})$$

$$=\frac{6+3\sqrt{7}+3\sqrt{5+2\sqrt{7}}}{\sqrt{12+6\sqrt{7}}\sqrt[4]{5+2\sqrt{7}}}.$$

The correctness of this section's algorithm hinges largely on the following lemma.

LEMMA 6. *Let $k\geq 1$ and $n\geq 0$. Let $r\in K_n$ and, for all $1\leq i\leq k$, let $a_i, b_i\in K_n$. Then*

$$\prod_{i=1}^{k}\sqrt{a_i+b_i\sqrt{r}}$$

denests using only reals if and only if at least one of the following holds:

(*i*) *$\sqrt{r}$ denests using only reals, or*

(*ii*) *$\prod_{i=1}^{k}\sqrt{a_i^2-b_i^2r}$ denests using only reals, or*

(*iii*) *$\sqrt{-r}\prod_{i=1}^{k}\sqrt{a_i^2-b_i^2r}$ denests using only reals.*

PROOF. Let $a, b\in K_n$ satisfy

$$a+b\sqrt{r}=\prod_{i=1}^{k}(a_i+b_i\sqrt{r}).$$

By Theorems 1–3, $\sqrt{a+b\sqrt{r}}$ denests using only reals if and only if at least one of the following holds:

(i) $\sqrt{r}$ denests using only reals, or

(ii) $\sqrt{a^2-b^2r}$ denests using only reals, or

(iii) $\sqrt{-r}\sqrt{a^2-b^2r}$ denests using only reals.

Hence it suffices to show that, under the hypothesis $\sqrt{r}\notin K_n$,

$$a^2-b^2r=\prod_{i=1}^{k}(a_i^2-b_i^2r).$$

This is done by a straightforward induction on k. □

The correctness of the algorithm follows by taking $m=1$ in the next theorem.

THEOREM 5. *Let $m\geq 1$ and $n\geq 1$. Given m formulas $\sqrt{f_1}, \sqrt{f_2}, \ldots, \sqrt{f_m}$, where $f_1, f_2, \ldots, f_m\in K_{n-1}$, procedure* DENEST *will correctly denest one of the following products in some real extension of Q, whenever possible:*

(*i*) *$\sqrt{f_m}$, or*

(*ii*) *$\prod_{i\in I}\sqrt{f_i}$, for some nonempty $I\subseteq\{1, 2, \ldots, m-1\}$, or*

(*iii*) *$\sqrt{-f_m}\prod_{i\in I}\sqrt{f_i}$, for some nonempty $I\subseteq\{1, 2, \ldots, m-1\}$.*

PROOF. The proof is by induction on n.

Basis. ($n=1$): This is precisely what the algorithm's basis does.

Induction ($n>1$): Let one of the products that denests be rewritten as

$$\sqrt{a+b\sqrt{r_{n-2}}}=\prod_{i\in I}\sqrt{c_i+d_i\sqrt{r_{n-2}}},$$

where $r_{n-2}\in K_{n-2}$, and $c_i, d_i\in K_{n-2}$ for all $i\in I$. By Lemma 6, this product denests using only reals if and only if at least one of the following products denests using only reals:

(i) $\sqrt{r_{n-2}}$, or

(ii) $\prod_{i\in I}\sqrt{c_i^2-d_i^2r_{n-2}}$, or (4.1)

(iii) $\sqrt{-r_{n-2}}\prod_{i\in I}\sqrt{c_i^2-d_i^2r_{n-2}}$.

Since $\sqrt{r_{n-2}}$ and $\sqrt{c_i^2-d_i^2r_{n-2}}$ (for all $i\in I$) are among the formulas passed in the recursive invocation, the induction hypothesis says some such product will be denested correctly in some real extension of Q; assume it is one of the three products in (4.1). If

$d=\sqrt{r_{n-2}}$ denested and is real, then surely

$$\sqrt{f_m}=\sqrt{a_m+b_m d}$$

is a correct denesting and is real. If

$$d=\prod_{i\in I}\sqrt{c_i^2-d_i^2 r_{n-2}}=\sqrt{\prod_{i\in I}(c_i^2-d_i^2 r_{n-2})}=\sqrt{a^2-b^2 r_{n-2}}$$

denested and is real, then the correctness of the denesting

$$\sqrt{a+b\sqrt{r_{n-2}}}=\frac{s+2b\sqrt{r_{n-2}}}{2\sqrt{s}},$$

where $s=2(a+d)$ is as in Theorem 4. Notice that $\sqrt{s}$ must be real, since all other subexpressions are real. Finally, if

$$d=\sqrt{-r_{n-2}}\prod_{i\in I}\sqrt{c_i^2-d_i^2 r_{n-2}}=\sqrt{-r_{n-2}(a^2-b^2 r_{n-2})}$$

denested, then the correctness of

$$\sqrt{a+b\sqrt{r_{n-2}}}=\frac{s+2a\sqrt{r_{n-2}}}{2\sqrt{s}\sqrt[4]{r_{n-2}}},$$

where $s=2(br_{n-2}+d)$ is most easily seen by squaring both sides. That $\sqrt{s}$ is real is proved as before. □

5. Denesting Rational Combinations of Radicals

So far we have developed the theory needed to denest certain single (possibly multiply nested) radicals. Suppose we have a linear combination of several radicals, none of which denests individually. Is it possible that the entire expression could denest? The answer is yes. For example, none of $\sqrt{1+\sqrt{3}}$, $\sqrt{3+3\sqrt{3}}$, and $\sqrt{10+6\sqrt{3}}$ denests, but

$$\sqrt{1+\sqrt{3}}+\sqrt{3+3\sqrt{3}}-\sqrt{10+6\sqrt{3}}$$

equals 0 and hence clearly denests.

Let K be a real extension of Q, and let $l_i, a_i, b_i, \sqrt{r_i}\in K$ for $1\leqslant i\leqslant h$, with $a_i+b_i\sqrt{r_i}>0$. (In fact, this last condition turns out to be unnecessary: although we need it in order to apply Lemma 1 in what follows, a weaker version of Lemma 1 involving only square roots could have been used.)

Suppose that the linear combination

$$\sum_{i=1}^{h} l_i\sqrt{a_i+b_i\sqrt{r_i}}$$

denests in K. Then it can be denested as follows. First denest any individual radical that denests. As we shall see, it suffices to determine, for each pair of remaining radicals $\sqrt{a_i+b_i\sqrt{r_i}}$ and $\sqrt{a_j+b_j\sqrt{r_j}}$, if their product denests in K. If the procedures of sections 2–4 show that $\sqrt{a_i+b_i\sqrt{r_i}}\sqrt{a_j+b_j\sqrt{r_j}}$ denests as $k\in K$, then replace $l_i\sqrt{a_i+b_i\sqrt{r_i}}+l_j\sqrt{a_j+b_j\sqrt{r_j}}$ by

$$\left(l_i+\frac{kl_j}{a_i+b_i\sqrt{r_i}}\right)\sqrt{a_i+b_i\sqrt{r_i}}\in K(\sqrt{a_i+b_i\sqrt{r_i}}),$$

and iterate the process of looking for a pair of radicals that denests.

If at some point no product of a pair of radicals denests, then we claim that the entire linear combination could not denest. For let $I=\{i_1, i_2, \ldots, i_m\}$ be a maximal set that satisfies

$$[K(\sqrt{a_{i_1}+b_{i_1}\sqrt{r_{i_1}}}, \ldots, \sqrt{a_{i_m}+b_{i_m}\sqrt{r_{i_m}}}):K]=2^m.$$

By Lemma 1, for every j in $\{1, 2, \ldots, h\}-I$,

$$\sqrt{a_j+b_j\sqrt{r_j}}=a\prod_{i\in I_j}\sqrt{a_i+b_i\sqrt{r_i}},$$

for some $I_j\subseteq I$ and some $a\in K$. But the 2^m distinct products are linearly independent, so the linear dependence can occur only if $I_j=I_k$ for some $j\neq k$, in which case $\sqrt{a_j+b_j\sqrt{r_j}}\sqrt{a_k+b_k\sqrt{r_k}}$ denests in K.

Note that linear combinations of radicals is as general as rational combinations, since a product of radicals is itself a radical, and quotients of radicals can be eliminated using conjugates.

Consider the example

$$\sqrt{1+\sqrt{3}}+\sqrt{3+3\sqrt{3}}-\sqrt{10+6\sqrt{3}}.$$

The product $\sqrt{1+\sqrt{3}}\sqrt{3+3\sqrt{3}}$ denests as $\sqrt{3}(1+\sqrt{3})$. Thus we simplify the expression to

$$\sqrt{1+\sqrt{3}}(1+\sqrt{3})-\sqrt{10+6\sqrt{3}}.$$

Next the product $\sqrt{1+\sqrt{3}}\sqrt{10+6\sqrt{3}}$ denests as $4+2\sqrt{3}$. Thus we can simplify further.

$$\begin{aligned}\sqrt{1+\sqrt{3}}+\sqrt{3+3\sqrt{3}}-\sqrt{10+6\sqrt{3}}&=\sqrt{1+\sqrt{3}}\left[1+\sqrt{3}-\frac{4+2\sqrt{3}}{1+\sqrt{3}}\right]\\&=\sqrt{1+\sqrt{3}}\left[\frac{(4+2\sqrt{3})-(4+2\sqrt{3})}{1+\sqrt{3}}\right]\\&=0.\end{aligned}$$

6. Denesting Arbitrary Formulas of Depth 2

In this section we consider a slightly different tower of fields

$$Q_0=Q,$$

and

$$\text{for } i\geqslant 0,\ Q_{i+1}=Q_i(\sqrt{r_i}),$$

where $r_i\in Q$ is positive. Using the procedure of section 3, we can assume that $\sqrt{r_i}\notin Q_i$ for all i. What makes the field Q_h different from the field K_h considered in section 4 is that any element of Q_h has depth at most 1 over Q, for all h.

Let $f\in Q_h$. In this section we describe an algorithm that denests $\sqrt{f}$ using square roots only, whenever such a denesting is possible. As a formula over Q, f is the linear combination of as many as 2^h different square roots, so the algorithm of section 2 is inapplicable to $\sqrt{f}$. As a formula over Q_{h-1}, however, f can be written as

$$f=a+b\sqrt{r_{h-1}},$$

where $a, b\in Q_{h-1}$. Thus, we can use the algorithm of section 2 to denest

$\sqrt{f} = \sqrt{a+b\sqrt{r_{h-1}}}$ over Q_{h-1}. (The reader is referred back to page 170 for a reminder of the definitions of "depth over Q_{h-1}" and "denest over Q_{h-1}".)

Since $\sqrt{r_{h-1}}$ is not an element of Q_{h-1}, $\sqrt{a+b\sqrt{r_{h-1}}}$ has depth 2 over Q_{h-1}. When the algorithm of section 2 denests $\sqrt{a+b\sqrt{r_{h-1}}}$ over Q_{h-1}, it will be rewritten as a formula of depth 1 over Q_{h-1}, that is

$$\sqrt{a+b\sqrt{r_{h-1}}} \in Q_{h-1}(\sqrt{r_{h-1}}, \sqrt{a_1}, \sqrt{a_2}, \ldots, \sqrt{a_k}),$$

for some $a_1, a_2, \ldots, a_k \in Q_{h-1}$. Notice that this "denested" version, like $\sqrt{a+b\sqrt{r_{h-1}}}$, still has depth 2 over Q, since $a_1, a_2, \ldots, a_k$ in general have depth 1 over Q. The progress we have made is that we have eliminated one radical, $\sqrt{r_{h-1}}$, from all depth 2 subformulas, since $a_1, a_2, \ldots, a_k$ are elements of Q_{h-1}, but $\sqrt{r_{h-1}}$ is not. Thus, we can denest $\sqrt{a_1}, \sqrt{a_2}, \ldots, \sqrt{a_k}$ recursively.

Given a positive $f \in Q_h$, the procedure that denests $\sqrt{f}$ over Q will return $d \in Q_h$ and positive $s \in Q$ satisfying $\sqrt{f} = d\sqrt{s}$. The procedure itself is as follows:

1. If $h = 0$, then return d if $f = d^2$ for some $d \in Q$, and fail otherwise. If $h > 0$, then write $f = a + b\sqrt{r_{h-1}}$, where $a, b \in Q_{h-1}$.
2. Compute $a^2 - b^2 r_{h-1} \in Q_{h-1}$.
3. Recursively denest $\sqrt{a^2 - b^2 r_{h-1}}$ over Q. If this fails to denest, then $\sqrt{f}$ fails to denest. Otherwise, the recrusive invocation returns $e \in Q_{h-1}$ and $s \in Q$ satisfying $\sqrt{a^2 - b^2 r_{h-1}} = e\sqrt{s}$.
4. Use the procedure of section 3 to determine whether $\sqrt{s} \in Q_{h-1}$. If not, $\sqrt{f}$ fails to denest. If so, the procedure expresses $\sqrt{s}$ as an element of Q_{h-1}.
5. Let $d = e\sqrt{s} \in Q_{h-1}$, and $t = 2(a+d) \in Q_{h-1}$. Then

$$\sqrt{a+b\sqrt{r_{h-1}}} = \frac{t + 2b\sqrt{r_{h-1}}}{2\sqrt{t}}.$$

   If $t \in Q$, then return

$$\frac{1}{2} + \frac{b\sqrt{r_{h-1}}}{t} \in Q_h \text{ and } t,$$

   and halt.
6. Recursively denest $\sqrt{t}$ over Q. If this fails to denest, then $\sqrt{f}$ fails to denest. Otherwise, the recursive invocation returns $j \in Q_{h-1}$ and $u \in Q$ satisfying $\sqrt{t} = j\sqrt{u}$.
7. $\sqrt{a+b\sqrt{r_{h-1}}} = \dfrac{t + 2b\sqrt{r_{h-1}}}{2j\sqrt{u}}$, so return $\dfrac{t + 2b\sqrt{r_{h-1}}}{2uj} \in Q_h$ and $u \in Q$.

Unlike all of the previous procedures, this one has exponential arithmetic complexity (measured as a function of h). Notice, though, that some formulas f in Q_h may have size 2^h; for these, the arithmetic complexity is polynomial in the length of f.

The correctness of the procedure is the subject of the next theorem.

THEOREM 6. *Given a positive $f \in Q_h$, the procedure will denest $\sqrt{f}$ over Q correctly whenever $\sqrt{f}$ can be denested using square roots only. Furthermore, when it succeeds $\sqrt{f}$ will be expressed as $\sqrt{f} = d\sqrt{s}$, for some $d \in Q_h$ and positive $s \in Q$.*

PROOF. The proof is by induction on h.

Basis ($h = 0$): $\sqrt{f}$ has depth 1 over Q, so $\sqrt{f}$ can be denested if and only if $f = d^2$ for some $d \in Q$. (In this case, $s = 1$ in the statement of the theorem.)

Induction ($h > 0$): Let $f = a + b\sqrt{r_{h-1}}$, where $a, b \in Q_{h-1}$ and $r_{h-1} \in Q$. Suppose $\sqrt{f}$ can be denested over Q using square roots only. Then certainly $\sqrt{f}$ can be denested over Q_{h-1} using square roots only, that is,

$$\sqrt{a+b\sqrt{r_{h-1}}} \in Q_{h-1}(\sqrt{r_{h-1}}, \sqrt{a_1}, \ldots, \sqrt{a_k}),$$

for some $a_1, a_2, \ldots, a_k \in Q_{h-1}$. Take $K = Q_{h-1}$ in Theorem 1, and recall that $\sqrt{r_{h-1}} \notin Q_{h-1}$. Hence, applying Theorem 1, $\sqrt{a^2 - b^2 r_{h-1}} \in Q_{h-1}$. Since $\sqrt{a^2 - b^2 r_{h-1}}$ has depth over Q one greater than any element in Q_{h-1}, $\sqrt{a^2 - b^2 r_{h-1}}$ denests over Q using square roots only. By the induction hypothesis, the recursive invocation in step (3) will denest this correctly as $\sqrt{a^2 - b^2 r_{h-1}} = e\sqrt{s}$, for some $e \in Q_{h-1}$ and positive $s \in Q$.

Since $e\sqrt{s} = \sqrt{a^2 - b^2 r_{h-1}} \in Q_{h-1}$ and $e \in Q_{h-1}$, $\sqrt{s} \in Q_{h-1}$. The procedure of section 3 will succeed in expressing $\sqrt{s}$ as an element of Q_{h-1}.

The correctness of

$$\sqrt{a+b\sqrt{r_{h-1}}} = \frac{t + 2b\sqrt{r_{h-1}}}{2\sqrt{t}}$$

in step (5) is as in Theorem 4. If $t \in Q$, then the right-hand side has depth 1 over Q, so $\sqrt{f}$ has been denested over Q. In this case, the only point to verify is that t is positive. But this follows from the fact that

$$\sqrt{t} = \frac{t + 2b\sqrt{r_{h-1}}}{2\sqrt{a+b\sqrt{r_{h-1}}}}$$

is real.

Assume, then, that $t \in Q_{h-1} - Q$, that is, $\sqrt{t}$ has depth 2 over Q, so the denesting is not completed. Note that

$$\sqrt{t} = \frac{t + 2b\sqrt{r_{h-1}}}{2\sqrt{a+b\sqrt{r_{h-1}}}}.$$

Since by hypothesis $\sqrt{a+b\sqrt{r_{h-1}}}$ can be denested over Q using square roots only, and since $t + 2b\sqrt{r_{h-1}} \in Q_h$ has depth 1 over Q, $\sqrt{t}$ can be denested over Q using square roots only. By the induction hypothesis, the recursive invocation of step (6) will return $j \in Q_{h-1}$ and positive $u \in Q$ satisfying $\sqrt{t} = j\sqrt{u}$. The correctness of step (7) is then simple to verify. □

Consider the example of denesting

$$\sqrt{a+b\sqrt{r_1}} = \sqrt{(112+70\sqrt{2})+(46+34\sqrt{2})\sqrt{5}}.$$

Here $h = 2$, $a = 112 + 70\sqrt{2}$, $b = 46 + 34\sqrt{2}$, and $r_1 = 5$. Furthermore, $Q_0 = Q$, $Q_1 = Q(\sqrt{2})$, and $Q_2 = Q(\sqrt{2}, \sqrt{5})$. The first task is to denest

$$\sqrt{a^2 - b^2 r_1} = \sqrt{(112+70\sqrt{2})^2 - (46+34\sqrt{2})^2 \times 5} = \sqrt{204+40\sqrt{2}}$$

recursively. This proceeds as in section 2, with the resulting denesting

$$d = \sqrt{204+40\sqrt{2}} = 2 + 10\sqrt{2}.$$

Step (4) would express d as an element of $Q_1 = Q(\sqrt{2})$, which it already is. In step (5), $t = 228 + 160\sqrt{2}$ and

$$\sqrt{a + b\sqrt{r_1}} = \frac{(228 + 160\sqrt{2}) + 2(46 + 34\sqrt{2})\sqrt{5}}{2\sqrt{228 + 160\sqrt{2}}}.$$

The numerator is now denested, but $\sqrt{t}$ in the denominator is not. However, $\sqrt{t}$ contains no occurrence of $\sqrt{5}$, so we proceed to denest $\sqrt{t}$ recursively in step (6). This proceeds again as in section 2, with the resulting denesting

$$\sqrt{t} = \sqrt{228 + 160\sqrt{2}} = 8\sqrt{2} + 10.$$

Finally, in step (7),

$$\sqrt{a + b\sqrt{r_1}} = \frac{(228 + 160\sqrt{2}) + 2(46 + 34\sqrt{2})\sqrt{5}}{2(8\sqrt{2} + 10)}$$

$$= (5 + 4\sqrt{2}) + (3 + \sqrt{2})\sqrt{5}.$$

Why does this approach not work if the input formula has depth $n > 2$ over Q? The problem is that $\sqrt{t}$ computed in step (5) in general has depth $n-1$ over Q_{h-1}, which in turn will have elements of depth $n-1$ over Q. Hence, $\sqrt{t}$ will have depth $2n-2$ over Q, so the formula's depth over Q actually *increases* if $n > 2$.

The authors are indepted to Clifford Earle and Oscar Rothans for the proof of Lemma 5.

References

Besicovitch, A. S. (1940). On the linear independence of fractional powers of integers. *J. London Math. Soc.* **15**, 3–6.

Caviness, B. F., Fateman, R. J. (1976). Simplification of radical expressions. *Proc. SYMSAC 77*. New York.

Herstein, I. N. (1975). *Topics in Algebra*. New York: John Wiley & Sons, Inc.

Richards, I. (1974). An application of Galois theory to elementary arithmetic. *Adv. Math.* **13**, 268–273.

Richardson, D. (1968). Some unsolved problems involving elementary functions of a real variable. *J. Symbolic Logic* **33**, 514–520.

Schinzel, A. (1982). *Selected Topics on Polynomials*. Michigan: The University of Michigan Press.

Zippel, R. E. B. (1977). *Simplification of Radicals with Applications to Solving Polynomial Equations*. M.Sc. thesis, Massachusetts Inst. of Technology.

J. Symb. Comp. 1 (1985), 189

Simplification of Expressions Involving Radicals

RICHARD ZIPPEL

Laboratory for Computer Science, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, *U.S.A.*

Algebraically dependent expressions arise in a large variety of symbolic computations. People seem to have the best intuition about expressions involving radicals. Symbolic computations with simple, non-nested, radicals is relatively straightforward; however, when the radicals are nested the problem becomes more difficult. This paper presents an algorithm for determining a linearly independent basis for a set of radicals (nested or not). This allows elementary techniques to be used for arithmetic operations on expressions involving elements of this set. In addition we provide a structure theorem that provides a sufficient condition for a nested radical to be expressed in terms of radicals of lower nesting level. These two techniques are powerful tools for computations involving radicals.

Algebraic manipulation is the study of techniques for performing mathematical computations symbolically. The most carefully studied and best understood problems in algebraic manipulation have been those that involve polynomials and rational functions—expressions involving independent variables combined by addition, multiplication and division. More complex expressions can often be treated as polynomials if the algebraically independent *kernels* can be identified. For instance, $e^{2x} + e^x$ can be treated as a polynomial in the kernel e^x.

It would be erroneous to view $e^{2x} + e^x$ as a polynomial in e^{2x} and e^x since they are not *independent*. They satisfy the polynomial relation $A - B^2$. Expressions that satisfy polynomial relations are *algebraically dependent*. Often more than one kernel is needed to express an expression as a polynomial over some simple ring. In this case the set of kernels used is called a *basis*. Ideally, the elements of the basis should be algebraically independent.

When algebraic numbers ($\sqrt[3]{7}$) or algebraic functions ($\sqrt{x^2-3}$) are involved things become a little more complex. Algebraic expressions cannot be used as kernels of polynomials since they must be reduced when the kernels appear to high enough powers ($\sqrt{2}^2 \to 2$). We are looking for independence results that ensure this is the only simplification we need to look for.

Regardless of whether the kernels are transcendental or algebraic, there are two basic problems that must be resolved when introducing new types of expressions to an algebraic manipulation system:

- —Determining a "sufficiently independent" basis for the computation, so that the expressions can be treated as polynomials.
- —Determining simple expressions for the basis elements so the user can understand the resulting expressions.

0747-7171/85/020189+22 $03.00/0